

## TANI YÖNTEMLERİNİN GÜVENİRLİĞİNİN VE GEÇERLİLİĞİNİN DEĞERLENDİRİLMESİ

Osman Günay\*

**Özet:** Bu yazıda, tanı yöntemlerinin güvenilirliği ve geçerliliği üzerinde duruldu. Güvenirliği ve geçerliliği etkileyen faktörler tartışıldı. Hassasiyet, seçicilik ve prediktif değerlerin hesaplanması anlatıldı.

**Anahtar kelimeler:** Tanı yöntemleri, güvenilirlik, geçerlilik

### GÜVENİRLİK (Reliability)

Tanı yönteminin güvenilirliği sonuçların tekrar edilebilirliğine (repeatability) bağlıdır. Tekrar edilebilirlik ise gözlem ve ölçümlerin tekrarlanması halinde aynı sonuçların bulunması anlamına gelir. Tekrar edilebilirliği etkileyen en önemli faktörler 1. gözlemcilerle, 2. araç-gereçle, 3. gözlenen olayla ilgili varyasyonlardır (1, 6).

1. Gözlemci varyasyonu (observer variation): Tanı yöntemini uygulayan kişi yada kişilerin aynı olayı farklı biçimde gözlemeleridir. Aynı gözlemci aynı olayı tekrar gözlediğinde farklı sonuçlar elde edebilir.

### Evaluating the reliability and validity of the diagnostic procedures

**Summary:** In this paper, the reliability and validity of the diagnostic procedures was explained and the factors affecting the reliability and validity were discussed. Calculations of sensitivity, specificity and predictive values were described.

**Key words:** Diagnostic procedures, reliability, validity

Örneğin, bir laborant aynı kan örneğinde iki kez hemoglobin ölçecek olursa, farklı sonuçlar elde edebilir. Bu tür varyasyon, gözlemci içi varyasyon (intra-observer variation) olarak isimlendirilir. Farklı gözlemcilerin aynı olayı farklı biçimde gözlemeleri ise gözlemciler arası varyasyona (inter-observer variation) neden olur. Örneğin, aynı kan örneğinde iki ayrı laborant farklı hemoglobin değerleri elde edebilir.

Gözlemci varyasyonunu tamamen ortadan kaldırmak mümkün olmamakla beraber, tanı yöntemlerini uygulayan kişilere, uyguladıkları yöntemler konusunda yeterli ve standart bir

---

\* Erciyes Üniversitesi Tıp Fakültesi Halk Sağlığı Anabilim Dalı Doçenti

eğitim verilerek ve sürekli denetlenerek, gözlemci varyasyonu azaltılabilir. Öte yandan, çeşitli istatistiksel yöntemler kullanılarak, gözlemci içi ve gözlemciler arası varyasyon kontrol edilebilir. Örneğin, laboratuvarında yeni çalışmaya başlayan bir elemana aynı ölçümler birkaç kez yaptırılarak, ölçümler arasındaki tutarsızlıklar belirlenebilir. Benzer şekilde, aynı ölçümler iki ayrı elemana ayrı ayrı yaptırılarak, gözlemciler arası tutarsızlıklar belirlenebilir. Bu tutarsızlıkların anlamlı olup olmadığı istatistiksel yöntemlerle test edilebilir.

2. Araç-gereçle ilgili varyasyon: Tanı yöntemlerinde kullanılan araç gerecin yeterince hassas ve standart olmamasına bağlı varyasyondur. Örneğin, kullanılan tansiyon aletlerinin standart olmaması, aynı kan basıncının farklı ölçülmesine neden olur. Bu nedenle tanı amacıyla kullanılan tüm ara-gereçlerin standardizasyonu sürekli olarak kontrol edilmelidir.

3. Gözlenen olayla ilgili varyasyon (biyolojik varyasyon): Tanı amacıyla gözlenen olayın farklı koşullarda farklı sonuçlar vermesi sonuçların tekrar edilebilirliğini olumsuz yönde etkiler. Örneğin, kan şekerinin yemekle, kan basıncının fizik aktivite ile ilişkili olması farklı koşullarda farklı sonuçların alınmasına neden olur. Bu nedenle, incelenecek değişkeni etkileyebilecek faktörler iyi bilinmeli ve gözlemler bu faktörler yönünden standart koşullarda yapılmalıdır. Örneğin, kan şekerinin yemekten belli bir süre sonra, kan basıncının ise yeterli istirahatten sonra ölçülmesi daha standart sonuçların elde edilmesini sağlar.

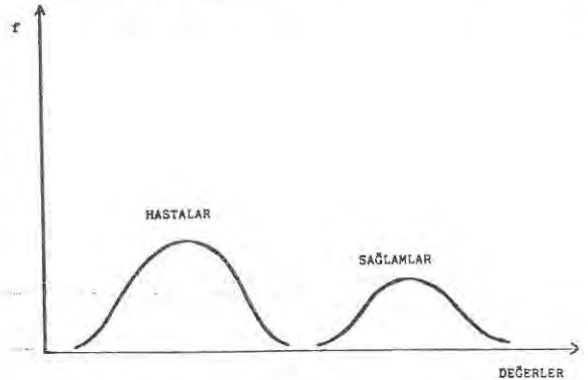
O halde, tanı yönteminin güvenilirliğini artırmak için, tüm ölçüm ve gözlemler yeterli ve standart eğitim görmüş personel tarafından, hassas ve standart araç-gereçle ve standart koşullarda yapılmalıdır. Aynı hastalığın tanısı için farklı tanı yöntemleri kullanılabilirse, güvenilirliği en yüksek olanlar tercih edilmelidir.

dir.

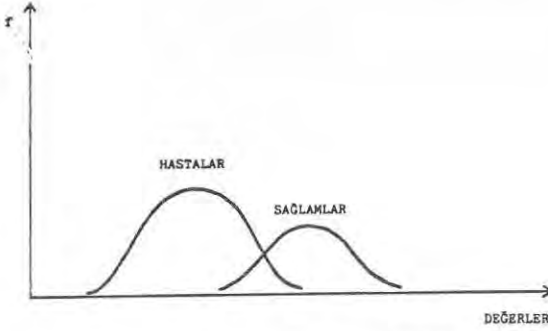
### GEÇERLİLİK (Validity)

Geçerlilik, tanı yönteminin hastalarla sağlamları birbirinden ayırma gücünü ifade eder (1, 6). Uygulanan tanı yönteminin güvenilirliği ne kadar yüksek olursa olsun, her zaman hasta ve sağlamları tam olarak ayıramaz. Bunun nedeni incelenen değişkenin hasta ve sağlamlardaki dağılımının tamamen birbirinden ayrı olmamasıdır. Örneğin, diabetik ve nondiabetik kişilerdeki açlık kan şekeri değerlerinin dağılımı birbirinden kesin olarak ayrılamaz. Açlık kan şekeri 95 mg/dl ölçülen bir kişi diabetik yada normal olabilir. Bu nedenle, açlık kan şekeri ne kadar doğru ölçülürse ölçülsün, diabetiklerle normal kişileri kesin olarak birbirinden ayıramaz.

Şekil 1'de hasta ve sağlamlara ait dağılımların birbirinden tamamen ayrı olduğu, şekil 2'de ise dağılımların kısmen çakıştığı hipotetik örnekler görülmektedir. Ne yazık ki, tanı amacıyla incelenen değişkenlerin dağılımı çoğunlukla şekil 2'dekine benzemektedir. Bu nedenle, uygulanan tanı yöntemlerinin hasta ve sağlamları ne ölçüde belirlediğinin test edilmesi zorunludur.



Şekil 1. Hasta ve Sağlamlara Ait Dağılımların Birbirinden Tamamen Ayrı Olduğu Hipotetik Örnek.



Şekil 2. Hasta ve Sağlamlara Ait Dağılımların Kısmen Çakıştığı Hipotetik Örnek.

Tanı yöntemlerinin geçerliliği, hassasiyeti (sensitivity) ve seçiciliği (specificity) ile ölçülür (1, 2, 3, 6, 7).

Hassasiyet; Tanı yönteminin hastaları belirleyebilme gücüdür.

Seçicilik; Tanı yönteminin sağlamları belirleyebilme gücüdür.

Hassasiyet ve seçiciliğin test edilebilmesi için, hasta ve sağlamlardan oluşan iki grupta bu tanı yönteminin uygulanması zorunludur. Bu amaçla kişilerin hasta olup olmadığını belirleyecek bir kesin tanı yöntemine (gold standart test) ihtiyaç vardır. Bu kesin tanı yöntemi

genellikle, uygulaması zor, pahalı ve invaziv bir yöntemdir (5,7). Öte yandan her hastalık için kesin tanı koydurucu bir tanı yöntemi yoktur. Bu nedenle otopsi sonuçlarından yararlanmak gerekebilir (2).

Tablo I. de görüldüğü gibi geçerliliği test edilecek tanı yöntemi tüm hastaların bir kısmını (a) doğru olarak belirlerken, bir kısmında (c) yalancı negatif sonuç vermiştir. Bu duruma göre yöntemin hassasiyeti şöyle bulunabilir;

$$\text{Hassasiyet} = \frac{\text{Doğru pozitifler}}{\text{Gerçek Hasta Sayısı}} \times 100 = \frac{a}{a+c} \times 100$$

Öte yandan, bu tanı yöntemi tüm sağlamların bir kısmını (d) doğru olarak belirlerken, bir kısmında (b) yalancı pozitif sonuç vermiştir. Bu duruma göre, yöntemin seçiciliği şöyle bulunabilir;

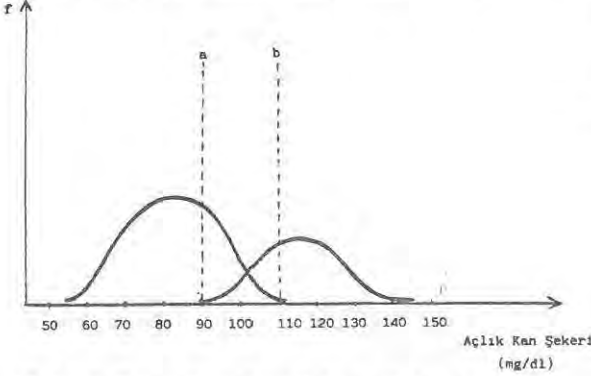
$$\text{Seçicilik} = \frac{\text{Doğru negatifler}}{\text{Gerçek Sağlam Sayısı}} \times 100 = \frac{d}{b+d} \times 100$$

Tanı yönteminin uygulanmasında, pozitif ve negatifler arasındaki sınır değiştirilerek, testin hassasiyeti veya seçiciliği artırılabilir. Ancak bu iksinin aynı anda yükseltilmesi mümkün

Tablo I. Geçerliliği Test edilecek Tanı Yönteminin Standart Tanı Yöntemi İle Karşılaştırılması

TESTEDİLECEK TANI YÖNTEMİ	STANDART TANI YÖNTEMİ		TOPLAM
	HASTA	SAĞLAM	
POZİTİF	DOĞRU POZİTİF (a)	YALANCI POZİTİF (b)	a + b
NEGATİF	YALANCI NEGATİF (c)	DOĞRU NEGATİF (d)	c + d
TOPLAM	GERÇEK HASTALAR (a+c)	GERÇEK SAĞLAMLAR (b+d)	a+b+c+d

değildir. Çünkü, hassasiyet ve seçicilikten birinin yükseltilmesi diğerinde azalmaya neden olur. Bu durum şekil 3'de gösterilmiştir.



Şekil 3. Diabetik ve Nondiabetik Kişilerde Açlık Kan Şekeri Değerlerinin Hipotetik Dağılımı.

Şekil III'deki hipotetik örnekte görüldüğü gibi, açlık kan şekeri için normal ve yüksek değerlerinsınırı 90 mg/dl olarak alınacak olursa, diabetik vakaların hemen hepsinde açlık kan şekeri yüksek kabul edileceği için, yöntemin hassasiyeti % 100'e yakın olacak; buna karşılık, birçok nondiabetik kişinin açlık kan şekeri yüksek sayıldığı için, yalancı pozitifler artacak, yani testin seçiciliği düşük bulunacaktır. Açlık kan şekeri için normal sınırınının 110 mg/dl'ye yükseltilmesi halinde ise, testin hassasiyeti düşecek, seçiciliği yükselecektir.

Tanı yöntemlerinin hassasiyeti ve seçiciliği, genel olarak hastalığın görülme sıklığı ile ilişkili değildir. Bu nedenle, bu grupta tespit edilen değerler diğer gruplar için de geçerlidir. Ancak, kişideki bazı başka hastalıklar veya özellikler yalancı pozitif ya da yalancı negatif sonuçlara neden olabilirler. Bu da hassasiyetin ve seçiciliğin değişmesine yol açabilir. Örneğin HIV enfeksiyonunun tanısı için uygulanan ELISA testinde, sıtma hastalığı ya da sık kan transfüzyonu yapılmış olması gibi durumlar yalancı pozitif sonuçlara neden olabilmek-

tedir (4). Bu nedenle tanı yöntemlerinin geçerliliği her toplum için yeniden değerlendirilmelidir.

#### PREDİKTİF DEĞERLER (Predictability)

Prediktif değerler tanı yönteminin verdiği sonuçların doğruluk oranını gösterir. Tanı yönteminin pozitif ve negatif sonuçları için ayrı prediktif değerler hesaplanır. Pozitif testin prediktif değeri tanı yönteminin pozitif sonuç verdiği kişilerin ne kadarının gerçekten hasta olduğunu, negatif testin prediktif değeri ise, yöntemin negatif sonuç verdiği kişilerin ne kadarının gerçekten sağlam olduğunu gösterir (4, 6). Tablo I'deki hipotetik verilerden yararlanılarak, prediktif değerler şöyle hesaplanabilir.

##### Pozitif Testin Prediktif Değeri

$$(PTPD) = a / (a+b) \times 100$$

##### Negatif Testin Prediktif Değeri

$$(NTPD) = d / (b+d) \times 100$$

Prediktif değerler, tanı yönteminin hassasiyeti ve seçiciliği ile ilgili olduğu gibi, incelenen grupta hastalığına görülme sıklığı ile de ilgilidir (1, 4, 6) ve seçiciliği sabit olan bir tanı yöntemi farklı gruplarda uygulandığında, hastalığın görülme sıklığı azaldıkça pozitif testin prediktif değeri azalır; buna karşılık, negatif testin prediktif değeri artar. Tablo II'de hastalık prevalansı % 1 olan bir grupta, hassasiyeti % 99, seçiciliği % 98 olan bir tanı yöntemi uygulandığında, prediktif değerlerin hesaplanması gösterilmiştir.

Tablo II'deki verilere göre;

$$PTPD = (99 / 297) \times 100 = \% 33.3$$

$$NTPD = (9702 / 9703) \times 100 = \% 99.9$$

Görüldüğü gibi, tanı yönteminin hassasiyeti ve seçiciliği çok yüksek olmasına rağmen, hastalık sıklığının çok az olması nedeniyle,

**Tablo II.** Hastalık Sıklığı % 1 olan 10.000 Kişilik Bir Grupta Hassasiyeti % 99 Seçiciliği % 98 Olan Bir Tanı Yöntemi İle Elde Edilecek Sonuçlar

TESTEDİLECEK TANI YÖNTEMİ	KESİN TANI YÖNTEMİ		TOPLAM
	HASTA	SAĞLAM	
POZİTİF	99 (a)	198 (b)	297 (a+b)
NEGATİF	1 (c)	9702 (d)	9703 (c+d)
<b>TOPLAM</b>	100 (a+c)	9900 (b+d)	10000 (a+b+c+d)

pozitif testin prediktif değeri çok düşük bulunmuştur. Yalancı pozitifler doğru pozitiflerden çok daha fazladır. Bu sonuçlara göre tedavi verilecek olursa, 1 hastaya karşılık, yaklaşık 2 kişi gereksiz yere tedavi görecektir. Bu nedenle, pozitif testin prediktif değerini yükseltmek için, uygulanması zor ve masraflı olan tanı yöntemleri, hastalık ihtimalinin daha yüksek olduğu seçilmiş gruplarda uygulanmalıdır. Pozitif testin prediktif değerinin düşük olduğu, yani yalancı pozitiflik oranının yüksek olduğu durumlarda, pozitif test sonucu veren kişilerde bir doğrulama testine başvurulmalıdır.

5. Sox HC: *Probability theory in the use of diagnostic tests. Annals of Internal Medicine* 104: 60-66, 1986.

6. Vaughan JP, Morrow RH (ed): *Manual of Epidemiology For District Health Management. World Health Organization, Geneva 1989, pp 71-88*

7. WHO: *Environmental and Occupational Epidemiology (Pre-Publication Copy). Geneva 1990, pp 97-105.*

#### Kaynaklar

1. Barker JP: *Practical Epidemiology*. 2. ed, Edinburg London and New York 1976, pp 14-29.

2. Dawson-Sounders B, Trapp RG: *Basic and Clinical Biostatistics (International Edition)*. Prentice-Hall International Inc, London 1990, pp 229-243.

3. Mausner JS, Kramer S: *Epidemiology: An Introductory Text*, WB Saunders Company, Philadelphia 1985, pp 214-237.

4. *Population Reports: AIDS-A Public Health Crisis. Series L, No 6, July-August 1986.*